PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/52, 9/12, 15/70, 1/21, C12P 21/00

A1

(11) International Publication Number: WO 98/14589

(43) International Publication Date: 9 April 1998 (09.04.98)

(21) International Application Number: PCT/EP97/05391

(22) International Filing Date: 1 October 1997 (01.10.97)

(30) Priority Data:
96115873.0 3 October 1996 (03.10.96) EP
(34) Countries for which the regional or international application was filed: DE et al.

(71) Applicant *(for all designated States except US)*: BOEHRINGER MANNHEIM GMBH [DE/DE]; D–68298 Mannheim (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: ANKENBAUER, Waltraud [DE/DE]; Oberanger 18, D–82377 Penzberg (DE). MARKAU, Ursula [DE/DE]; Hungerwiese 1, D–82398 Polling (DE). SVETLICHNY, Vitaly [RU/DE]; Warmensteinacher Strasse 91G, D–95448 Bayreuth (DE). SCHMITZ–AGHEGUIAN, Gudrun [DE/DE]; Wettersteinstrasse 3, D–82347 Bernried (DE). REISER, Astrid [DE/DE]; Schleier–Weg 20, D–82387 Antdorf (DE). ANGERER, Bernhard [DE/DE]; Schützenstrasse 20, D–83024 Rosenheim (DE). EBENBICHLER, Christine [DE/DE]; Dorfstrasse 5, D–82387 Antdorf (DE). LAUE, Frank [DE/DE]; Seestrasse 16, D–82396 Paehl (DE). BONCH–OSMOLOVSKAYA, Elizaveta [RU/RU]; Lomonosovski AV, 33–2–60, Moscow, 117192 (RU).

(74) Common Representative: BOEHRINGER MANNHEIM GMBH; Patent Dept., D–68298 Mannheim (DE).

(81) Designated States: AU, CA, IL, JP, NO, NZ, RU, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: THERMOSTABLE DNA POLYMERASE FROM CARBOXYDOTHERMUS HYDROGENOFORMANS

(57) Abstract

A DNA polymerase from a thermophilic eubacterium is provided. The DNA polymerase shows magnesium ion dependent reverse transcriptase activity and 3'–5' exonuclease activity. The invention also includes recombinant plasmids and transformed host cells capable of producing the enzyme. The enzyme is classified into class EC 2.7.7.7., a DNA nucleotidyl transferase DNA–directed type.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Thermostable DNA Polymerase from Carboxydothermus hydrogenoformans

The present invention relates to a thermostable enzyme which is a DNA polymerase obtainable from Carboxydothermus hydrogenoformans.

Furthermore, the present invention relates to the field of molecular biology and provides improved methods for the replication and amplification of deoxyribonucleic (DNA) and ribonucleic acid (RNA) sequences. In a preferred embodiment, the invention provides a method for synthesizing a complementary DNA copy from an RNA template with a thermoreactive DNA polymerase. In another aspect, the invention provides methods for amplifying a DNA segment from an RNA or DNA template using a thermostable DNA polymerase (RT-PCR or PCR).

Heat stable DNA polymerases (EC 2.7.7.7. DNA nucleotidyltransferase, DNA-directed) have been isolated from numerous thermophilic organisms (for example: Kaledin et al. (1980), Biokimiya 45, 644-651; Kaledin et al. (1981) Biokimiya 46, 1247-1254; Kaledin et al. (1982) Biokimiya 47, 1515-1521; Ruttimann et al. (1985) Eur. J. Biochem. 149, 41-46; Neuner et al. (1990) Arch. Microbiol. 153, 205-207).

For some organisms, the polymerase gene has been cloned and expressed (Lawyer et al. (1989) J. Biol. Chem. 264, 6427-6437; Engelke et al. (1990) Anal. Biochem. 191, 396-400; Lundberg et al. (1991) Gene 108, 1-6; Perler et al. (1992) Proc. Natl. Acad. Sci. USA 89, 5577).

Thermophilic DNA polymerases are increasingly becoming important tools for use in molecular biology and there is growing interest in finding new polymerases which have more suitable properties and activities for use in diagnostic detection of RNA and DNA, gene cloning and DNA sequencing. At present, the thermophilic DNA polymerases

mostly used for these purposes are from Thermus species like Taq polymerase from T. aquaticus (Brock et al. (1969) J. Bacteriol. 98, 289-297)

Reverse transcription is commonly performed with viral reverse transcriptases like the enzymes isolated from Avian myeloblastosis virus or Moloney murine leukemia virus, which are active in the presence of Magnesium ions but have the disadvantages to possess RNase H-activity, which destroys the template RNA during the reverse transcription reaction and have a temperature optimum at at 42°C or 37 °C, respectively.

Alternative methodes are described using the reverse transcriptase activity of DNA polymerases of thermophilic organisms which are active at higher temperatures. Reverse transcription at higher temperatures is of advantage to overcome secondary structures of the RNA template which could result in premature termination of products. Thermostable DNA polymerases with reverse transcriptase activities are commonly isolated from Thermus species. These DNA polymerases however, show reverse transcriptase activity only in the presence of Manganese ions. These reaction conditions are suboptimal, because in the presence of Manganese ions the polymerase copies the template RNA with low fidelity.

Another feature of the commonly used reverse transcriptases is that they do not contain 3'-5'exonuclease activity. Therefore, misincorporated nucleotides cannot be removed and thus the cDNA copies from the template RNA may contain a significant degree of mutations.

Therefore, it is desirable to develop a reverse transcriptase

- which acts at higher temperatures to overcome secondary structures in the template to avoid premature termination of the reaction and to assure the production of cDNA without deletions
- which is active in the presence of Magnesium ions in order to prepare cDNA from RNA templates with higher fidelity and

- which has 3′-5′-exonuclease in order to remove misincorporated nucleotides before continuation of DNA synthesis and to produce a product with a low mutation frequency.

The present invention adresses these needs and provides a heat stable DNA polymerase active at higher temperatures which has reverse transcriptase activity in the presence of magnesium ions and and which has 3′-5′exonuclease activity.

It is an object of this invention to provide a polymerase enzyme (EC 2.7.7.7.), characterised in that it has reverse transcriptase activity in the presence of magnesium ions as well as in the presence of manganese ions. In another aspect the invention comprises a DNA polymerase isolated from Carboxydothermus hydrogenoformans (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, D-38124 Braunschweig, DSM No. 8979). In a further aspect the invention comprises a DNA polymerase having reverse transcriptase activity in the presence of magnesiums ions and in the substantial absence of manganese ions. In a further aspect the invention comprises a DNA polymerase having a molecular mass of about 100 to 105 kDa as determined by in situ PAGE analysis. In a further aspect the invention comprises a reverse transcriptase which is thermostable. In a further aspect the invention comprises a DNA polymerase having 3′-5′-exonuclease activity. In a further aspect the invention comprises a recombinant DNA sequence that encodes DNA polymerase activity of the microorganism Carboxidothermus hydrogenoformans. In a related aspect, the DNA sequence is depicted as SEQ ID No. 7. In a second related aspect the invention comprises a recombinant DNA sequence that encodes essentially amino acid residues 1 to 831. In a further aspect the invention comprises a recombinant DNA plasmid that comprises the DNA sequence of the invention inserted into plasmid vectors and which can be used to drive the expression of the thermostable DNA polymerase of Carboxydothermus hydrogenoformans in a host cell transformed with the plasmid. In a further aspect the invention includes a recombinant strain comprising the vector pDS56 carrying the Carboxydothermus hydrogenoformans DNA polymerase gene and designated pAR 4. The E.coli strain (BL21 (DE3)pUBS520) carrying the plasmid pAR4

was deposited on the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, D-38124 Braunschweig DSM No. 11179) and is designated AR96

In referring to a peptide chain as being comprised of a series of amino acids "substantially or effectively" in accordance with a list offering no alternatives within itself, we include within that reference any versions of the peptide chain bearing substitutions made to one or more amino acids in such a way that the overall structure and the overall function of the protein composed of that peptide chain is substantially the same as - or undetectably different to - that of the unsubstituted version. For example it is generally possible to exchange alanine and valine without greatly changing the properties of the protein, especially if the changed site or sites are at positions not critical to the morphology of the folded protein.

The DNA polymerase is "thermostable" meaning that it is stable to heat and preferentially active at higher temperatures, especially the high temperatures used for denaturation of DNA strands. More particularly, the thermostable DNA polymerases are not substantially inactivated at the high temperatures used in polymerase chain reactions.

The term "reverse transcriptase" describes a class of polymerases characterized as RNA-dependent DNA polymerases. All known reverse transcriptases require a primer to synthesize a DNA transcript from an RNA template. Historically, reverse transcriptase has been used primarily to transcribe mRNA into cDNA which can then be cloned into a vector for further manipulation.

Other definitions are used in a manner consistent with the art.

Carboxydothermus hydrogenoformans was isolated from a hot spring in Kamchatka by V. Svetlichny. A sample of C. hydrogenoformans was deposited on the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM) under the terms of the Budapest Treaty and received Accession Number DSM 8979. The thermostable

polymerase isolated from Carboxydothermus hydrogenoformans has a molecular weight of 100 to 105 KDa and retains more than 60 % of its initial activity after heating to 95°C for 30 minutes.The thermostable enzyme possesses a 5'-3' polymerase activity, a 3'-5'-exonuclease activity, a 5'-3'- exonuclease activity and a reverse transcriptase-activity which is Mg^{++}-dependent. The polymerase according to the present invention has reverse transcriptase activity in the presence of magnesium ions and in the substantial absence of manganese ions. The thermostable enzyme may be native or recombinant and may be used for first- and second-strand cDNA synthesis, in cDNA cloning, DNA sequencing, DNA labeling and DNA amplification.

For recovering the native protein C.hydrogenoformans may be grown using any suitable technique, such as the technique described by Svetlichny et al. (1991) System. Appl. Microbiol., 14, 205-208. After cell growth one preferred method for isolation and purification of the enzyme is accomplished using the multi-step process as follows:

The cells are thawed, suspended in buffer A (40 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 7 mM 2-mercaptoethanol, 0.4 M NaCl, 10 mM Pefabloc) and lysed by twofold passage through a Gaulin homogenizer. The raw extract is cleared by centrifugation, the supernatant dialyzed against buffer B (40 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 7 mM 2-mercaptoethanol, 10 % Glycerol) and brought onto a column filled with Heparin-Sepharose (Pharmacia). In each case the columns are equilibrated with the starting solvent and after the application of the sample washed with the threefold of its volume with this solvent. Elution of the first column is performed with a linear gradient of 0 to 0.5 M NaCl in Buffer B. The fractions showing polymerase activity are pooled and ammonium sulfate is added to a final concentration of 20 %. This solution is applied to a hydrophobic column containing Butyl-TSK-Toyopearl (TosoHaas). This time the column is eluted with a falling gradient of 20 to 0 % ammonium sulfate. The pool containing the activity is dialysed and again transferred to a column, this time with DEAE-Sepharose (Pharmacia), and eluted with a linear gradient of 0-0.5 M NaCl in buffer B. The fourth column contains Tris-Acryl-Blue (Biosepra) and is eluted as in the preceding case.

Finally the active fractions are dialyzed against buffer C (20 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 7.0 mM 2-mercaptoethanol, 100 mM NaCl, 50 % Glycerol.

Isolation of recombinant DNA polymerase from Carboxydothermus hydrogenoformans may be performed with the same protocol or with other commonly used procedures.

DNA polymerase activity was measured by incorporation of digoxigenin-labeled dUTP into the sythesized DNA and detection and quantification of the incorporated digoxigenin essentially according to the method described in Höltke, H.-J.; Sagner, G; Kessler, C. and Schmitz, G. (1992) Biotechniques 12, 104 -113.

Determination of reverse transcriptase activity is performed essentially as described for determination of DNA polymerase activity except that the reaction mixture consists of the components as described by example 3.

In situ PAGE analysis of polymerase activity and reverse transcriptase activity was performed essentially according to the method described by Spanos A. and Hübscher U. ((1983) Methods in Enzymology 91, 263-277). Some minor, but essential modifications to the original method are, that the renaturation of the SDS-denatured polypeptides is performed in the presence of magnesium ions (3 mM) and dATP (0.5-1 µM) to assist refolding.

3'-5' exonuclease activity is commonly referred as "proofreading" or "editing" activity of DNA polymerases . It is located in the small domain of the large fragment of Type A polymerases. This activity removes mispaired nucleotides from the 3' end of the primer terminus of DNA in the absence of nucleoside triphosphates (Kornberg A. and Baker T.A.(1992) DNA Replication W. H. Freemann & Company, New York). This nuclease action is suppressed by deoxynucleoside triphosphates if they match to the template and can be incorporated into the polymer.

The 3'- 5' exonuclease activity of the claimed DNA polymerase can be measured as degradation or shortening of a 5'-digoxygenin-labeled oligonucleotide annealed to

template DNA in the absence or presence of deoxyribonucleoside triphosphates or on DNA fragments in the absence or presence of deoxyribonucleoside triphosphates.

Carboxydothermus hydrogenoformans DNA polymerase is the first DNA polymerase isolated from thermophilic eubacteria with a higher activity in the presence of magnesium ions than in the presence of manganese ions as shown in figure 1. Compaired to the DNA polymerase activity the reverse transcriptase activity in the presence of magnesium is relatively high. This is shown - in comparison with DNA polymerases from T.filiformis, A.thermophilum and the most commonly used DNA polymerase for reverse transcription T.thermophilus in figure 6. The reverse transcriptase activity in dependence of magnesium is of advantage since the DNA polymerases synthesize DNA with higher fidelity in the presence of magnesium than in the presence of manganese (Beckmann R. A. et al. (1985) Biochemistry 24, 5810-5817; Ricchetti M. and Buc H. (1993) EMBO J. 12, 387-396). Low fidelity DNA synthesis is likely to lead to mutated copies of the original template. In addition, Mn^{2+} ions have been implicated in an increased rate of RNA degradation, particularly at higher temperatures and this can cause the synthesis of shortened products in the reverse transcription reaction.

The DNA sequence (SEQ ID No.: 7) of Carboxydothermus hydrogenoformans polymerase and the derived amino acid sequence of the enzyme are shown in figure 5. The molecular weight deduced from the sequence is 94 348 Da, in SDS polyacrylamide gel electrophoresis however the Carboxydothermus hydrogenoformans polymerase has an electrophoretic mobility higher than E.coli pol I (109 kDa) and a lower mobility than Taq polymerase (94 kDa) and Klenow fragment (76 kDa) as shown in figure 2. Comparing the migration properties of Taq and E.coli DNA polymerases with those of Carboxydothermus hydrogenoformans polymerase a molecular weight of 100 to 105 kDa can be deduced. Since the Carboxydothermus hydrogenoformans polymerase isolated from the native strain has the same migration properties as the recombinant enzyme the „slower" migration during SDS gel electrophoresis must rather be a property of the enzyme than a cloning artefact. A possible explanation for this phenomenon could be that

this enzyme which is derived from a thermophilic organism has a very stable structure which is not completely unfolded under the standard denaturation conditions used.

The production of a recombinant form of Carboxydothermus hydrogenoformans DNA polymerase generally includes the following steps: chromosomal DNA from Carboxydothermus hydrogenoformans is isolated by treating the cells with detergent e.g. SDS and a proteinase e.g. Proteinase K. The solution is extracted with phenol and chloroform and the DNA purified by precipitation with ethanol. The DNA is dissolved in Tris/ EDTA buffer and the gene encoding the DNA polymerase is specifically amplified by the PCR technique using two mixed oligonucleotides (primer 1 and 2). These oligonucleotides, described by SEQ ID No.: 1 and SEQ ID No.: 2, were designed on the basis of conserved regions of family A DNA polymerases as published by Braithwaite D. K. and Ito J., 1993, Nucl. Acids Res. Vol. 21, p. 787 - 802. The specifically amplified fragment is ligated into an vector, preferably the pCR™II vector (Invitrogen) and the sequence is determined by cycle-sequencing. Complete isolation of the coding region and the flanking sequences of the DNA polymerase gene can be performed by restriction fragmentation of the Carboxydothermus hydrogenoformans DNA with another restriction enzyme as in the first round of screening and by inverse PCR (Innis et al., (1990) PCR Protocols; Academic Press, Inc., p. 219-227). This can be accomplished with synthesized oligonucleotide primers binding at the outer DNA sequences of the gene part but in opposite orientation. These oligonucleotides described by SEQ ID Nos. 3 and 4 , were designed on the basis of the sequences which were determined by sequencing of the first PCR product described above. As template Carboxydothermus hydrogenoformans DNA is used which is cleaved by restriction digestion and circularized by contacting with T4 DNA ligase. To isolate the coding region of the whole polymerase gene, another PCR is performed using primers as shown in SEQ ID Nos. 5 and 6 to amplify the complete DNA polymerase gene directly from genomic DNA and introducing ends compatible with the linearized expression vector.

SEQ ID No. 1:

Primer 1: 5′-CCN AAY YTN CAR AAY ATH-3′

SEQ ID No. 2:

Primer 2: 5′-YTC RTC RTG NAC YTG-3′

SEQ ID No. 3:

Primer 3: 5'-GGG CGA AGA CGC TAT ATT CCT GAG C-3'

SEQ ID NO. 4:

Primer 4: 5'-GAA GCC TTA ATT CAA TCT GGG AAT AAT C-3'

SEQ ID NO. 5:

Primer 5: 5'-CGA ATT CAA TCC ATG GGA AAA GTA GTC CTG GTG GAT-3'

SEQ ID NO. 6:

Primer 6: 5'-CGA ATT CAA GGA TCC TTA CTT CGC TTC ATA CCA GTT-3'

The gene is operably linked to appropriate control sequences for expression in either prokaryotic or eucaryotic host/vector systems. The vector preferably encodes all functions required for transformation and maintenance in a suitable host, and may encode selectable markers and/or control sequences for polymerase expression. Active recombinant thermostable polymerase can be produced by transformed host cultures either continuously or after induction of expression. Active thermostable polymerase can be

recovered either from host cells or from the culture media if the protein is secreted through the cell membrane.

It is also preferable that Carboxydothermus hydrogenoformans thermostable polymerase expression is tightly controlled in E.coli during cloning and expression. Vectors useful in practicing the present invention should provide varying degrees of controlled expression of Carboxydothermus hydrogenoformans polymerase by providing some or all of the following control features: (1) promoters or sites of initiation of transcription, either directly adjacent to the start of the polymerase gene or as fusion proteins, (2) operators which could be used to turn gene expression on or off, (3) ribosome binding sites for improved translation, and (4) transcription or translation termination sites for improved stability. Appropriate vectors used in cloning and expression of Carboxydothermus hydrogenoformans polymerase include, for example, phage and plasmids. Example of phage include lambda gtII (Promega), lambda Dash (Stratagene) lambda ZapII (Stratagene). Examples of plasmids include pBR322, pBTac2 (Boehringer Mannheim), pBluescript (Stratagene), pSP73 (Promega), pET3A (Rosenberg, supra, (1987) Gene 56:125-135), pASK75 (Biometra), pDS56 (Stüber, D., Matile, H. and Garotta G. (1990) Immunological Methods, Letkovcs, I. and Pernis, B., eds.) and pET11C (Studier, F. W. (1990) Methods in Enzymology, 185:60-89). According to the present invention the use of a plasmid has shown to be advantageously, particularly pDS56. The Plasmid pDS56 carrying the Carboxydothermus hydrogenoformans DNA polymerase gene is then designated pAR4.

Standard protocols exist for transformation, phage infection and cell culture (Maniatis, et al. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press). Of the numerous E.coli strains which can be used for plasmid transformation, the preferred strains include JM110 (ATCC 47013), LE392 pUBS 520 (Maniatis et al. supra; Brinkmann et al., (1989) Gene 85:109-114;), JM101 (ATCC No. 33876), XL1 (Stratagene), and RR1 (ATCC no. 31343), BL21 (DE3) pUBS520 (Brinkmann, U. et al. (1989) Gene 85, 109-114) and BL21 (DE3) plysS (Studier, F. W. et al., (1990) Methods in Enzymology, supra). According to the present invention the use of the E. coli strain

BL21 (DE3) pUBS520 has shown to be advantageously. The E. coli strain BL21 (DE3) pUBS520 transformed with the plasmid pAR4 is then designated AR96(DSM No. 11179). E.coli strains XL1-Blue (Stratagene), and ER1458 (Raleigh, E. A. et al., (1988) Nucleic Acids Research 16:1563-1575) are among the strains that can be used for lambda phage, and Y1089 can be used for lambda gt11 lysogeny. The transformed cells are preferably grown at 37°C and expression of the cloned gene is induced with IPTG.

Isolation of the recombinant DNA polymerase can be performed by standard techniques. Separation and purification of the DNA polymerase from the E.coli extract can be performed by standard methods. These methods include, for example, methods utilizing solubility such as salt precipitation and solvent precipitation, methods utilizing the difference in molecular weight such as dialysis, ultra-filtration, gel-filtration, and SDS-polyacrylamide gel electrophoresis, methods utilizing a difference in electric charge such as ion-exchange column chromatography, methods utilizing specific interaction such as affinity chromatography, methods utilizing a difference in hydrophobicity such as reversed-phase high performance liquid chromatography and methods utilizing a difference in isoelectric point such as isoelectric focussing electrophoresis.

The present invention provides improved methods for efficiently transcribing RNA and amplifying RNA or DNA. These improvements are achieved by the discovery and application of previously unknown properties of thermoactive DNA polymerases.

The thermostable enzyme of this invention may be used for any purpose in which such enzyme activity is necessary or desired. In a particularly preferred embodiment, the enzyme catalyzes the nucleic acid amplification reaction known as PCR. This process for amplifying nucleic acid sequences is disclosed and claimed in U.S. Patent No. 4,683,202. The PCR nucleic acid amplification method involves amplifying at least one specific nucleic acid sequence contained in a nucleic acid or a mixture of nucleic acids and produces double-stranded DNA. Any nucleic acid sequence, in purified or nonpurified form, can be utilized as the starting nucleic acid(s), provided it contains or is sus-

pected to contain the specific nucleic acid sequence desired. The nucleic acid to be amplified can be obtained from any source, for example, from plasmids such as pBR322, from cloned DNA or RNA, from natural DNA or RNA from any source, including bacteria, yeast, viruses, organelles, and higher organisms such as plants and animals, or from preparations of nucleic acids made in vitro.

DNA or RNA may be extracted from blood, tissue material such as chorionic villi, or amniotic cells by a variety of techniques. See, e.g., Maniatis et al., 1982, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York) pp. 280-281. Thus the process may employ, for example, DNA or RNA, including messenger RNA, which DNA or RNA may be single-stranded or double-stranded. In addition, a DNA-RNA hybrid which contains one strand of each may be utilized.

The amplification of target sequences in DNA or from RNA may be performed to proof the presence of a particular sequence in the sample of nucleic acid to be analyzed or to clone a specific gene. DNA polymerase from Carboxydothermus hydrogenoformans is very useful for these processes. Due to its 3'-5'exonuclease activity it is able to synthesize products with higher accuracy as the reverse transcriptases of the state of the art.

DNA polymerase from Carboxydothermus hydrogenoformans may also be used to simplify and improve methods for detection of RNA target molecules in a sample. In these methods DNA polymerase from Carboxydothermus hydrogenoformans may catalyze: (a) reverse transcription, (b) second strand cDNA synthesis, and, if desired, (c) amplification by PCR. The use of DNA polymerase from Carboxydothermus hydrogenoformans in the described methods would eliminate the previous requirement of two sets of incubation conditions which are necessary due to the use of different enzymes for each step. The use of DNA polymerase from Carboxydothermus hydrogenoformans may be used to perform RNA reverse transcription and amplification of the resulting complementary DNA with enhanced specificity and with fewer steps than previous RNA cloning and diagnostic methods.

Brief description of the drawings

Fig. 1 shows the relative reverse transcriptase activity of DNA polymerase from Carboxydothermus hydrogenoformans in dependence of magnesium and manganese salt.

Fig. 2 shows a photograph of a DNA polymerase assay performed in situ. The activity of DNA polymerase from Carboxydothermus hydrogenoformans and reference polymerases is detected in situ. A fraction of DNA polymerase from C. hydrogenoformans and reference enzyme were submitted to electrophoresis on a SDS-polyacrylamide gel containing activated calf thymus DNA. After electrophoresis the SDS was removed, the proteins were renatured and incubated at 65°C in the presence of magnesium salt, dNTPs and digoxigenin labeled dUTP to allow DNA synthesis. The nucleic acid was blotted to a nylon membrane and the newly synthesized DNA detected by a chemiluminescence reaction.

Fig. 3 shows the thermostability of DNA polymerase from Carboxydothermus hydrogenoformans. Aliquots of the DNA polymerase were incubated for 30 min. at the temperatures indicated in the figure, and subsequently the remaining enzyme activity was determined.

Fig. 4 shows the analysis for 3´-5´-exonuclease activity of DNA polymerase from Carboxydothermus hydrogenoformans in comparison with DNA polymerase from Thermus aquaticus and Pyrococcus woeseii. The analysis is performed in the presence or absence of dNTPs. A 22mer primer labeled with digoxigenin at the 5´-end was annealed to a 34mer template DNA leaving a 12 bp 5´overhang of template DNA. DNA polymerases from Carboxydothermus hydrogenoformans, Thermus aquaticus and Pyrococcus woeseii were incubated with this substrate in the presence of magnesium with or without dNTPs. The products were separated on a sequencing gel, blotted to a nylon membrane and detected by a chemiluminescence reaction.

Fig. 5 shows the DNA sequence of the polymerase gene of Carboxydothermus hydrogenoformans with SEQ ID No.: 7 and the derived peptide sequence of the DNA polymerase protein with SEQ ID No. 8.

Fig. 6 shows a comparison of the reverse transcriptase activity thermostable DNA polymerases Carboxydothermus hydrogenoformans, Anaerocellum thermophilum, Thermus filiformis (Pacific Enzymes) and Thermus thermophilus. Similar amounts (units) of the DNA polymerases were analyzed. Each enzyme was tested for DNA polymerase activity, for reverse tanscriptase activity in the presence of Mg^{++} (5 mM) and reverse transcriptase activity in the presence of Mn^{++} (1 mM) under the reaction conditions optimal for the individual enzymes. DNA synthesis was measured by incorporation of digoxigenin-labeled nucleotides. In order to compare the ratio of DNA polymerase to reverse transcriptase activity, the relative light units (RLU) measured in the DNA polymerase assay was set to 100. The RLUs measured in the reverse transcriptase activity tests are expressed as percent of the polymerase activity.

Example 1

Detection of endonuclease, exonuclease and ribonuclease activities:

Absence of endonuclease activity: 1 µg of plasmid DNA is incubated for 4 hours with an excess of purified DNA polymerase in 50 µl of test buffer with a paraffin oil overlay at 72°C.

Absence of nonspecific exonuclease activity: 1 µg of EcoRI/HindIII-fragments of lambda DNA are incubated in 100 µl of test buffer in the absence and presence of dNTPs (1mM final concentration each) with an excess of purified DNA polymerase for 4 hours at 72°C.

Absence of ribonuclease activity: 3 µg of MS2 RNA are incubated with an excess of DNA polymerase in 20 µl of test buffer for 4 hours at 72°C. The RNA is subsequently

analyzed by electrophoresis in a MOPS gel (Maniatis et al. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, New York).

Example 2

Determination of DNA polymerase activity

DNA polymerase activity was measured by incorporation of digoxigenin-labeled dUTP into the sythesized DNA and detection and quantification of the incorporated digoxigenin essentially according to the method described in Höltke, H.-J.; Sagner, G; Kessler, C. and Schmitz, G. (1992) Biotechniques 12, 104 -113. The reaction is performed in a reaction volume of 50 µl containing 1 or 2 µl of diluted (0.05 U - 0.01 U) DNA polymerase and 50 mM Tris-HCl, pH 8.5; 12.5 mM $(NH_4)_2SO_4$; 10 mM KCl; 5 mM $MgCl_2$; 10 mM 2-mercaptoethanol; 33 µM dNTPs; 200 µg/ml BSA;12 µg of DNAse I-activated DNA from calf thymus and 0.036 µM digoxigenin-dUTP.

The samples are incubated for 30 min. at 72°C, the reaction is stopped by addition of 2 µl 0.5 M EDTA, and the tubes placed on ice. After addition of 8 µl 5 M NaCl and 150 µl of Ethanol (precooled to -20°C) the DNA is precipitated by incubation for 15 min. on ice and pelleted by centrifugation for 10 min at 13000 x rpm and 4°C. The pellet is washed with 100 µl of 70% Ethanol (precooled to -20°C) and 0.2 M NaCl, centrifuged again and dried under vacuum.

The pellets are dissolved in 50 µl Tris-EDTA (10 mM/0.1 mM; pH 7.5). 5 µl of the sample are spotted into a well of a nylon membrane bottomed white microwell plate (Pall Filtrationstechnik GmbH, Dreieich, FRG, product no: SM045BWP). The DNA is fixed to the membrane by baking for 10 min. at 70°C. The DNA loaded wells are filled with 100 µl of 0.45 µm-filtrated 1 % blocking solution (100 mM maleic acid, 150 mM NaCl, 1 % (w/v) casein, pH 7.5). All following incubation steps are done at room temperature. After incubation for 2 min. the solution is sucked through the membrane with

a suitable vacuum manifold at -0.4 bar. After repeating the washing step, the wells are filled with 100 µl of a 1:10 000-dilution of Anti-digoxigenin-AP, Fab fragments (Boehringer Mannheim, FRG, no:1093274) diluted in the above blocking solution. After incubation for 2 min. and sucking this step is repeated once. The wells are washed twice under vacuum with 200 µl each time washing-buffer 1 (100 mM maleic-acid, 150 mM NaCl, 0.3 %(v/v) Tween™ 20 (Poly(oxyethylen)$_n$-Sorbitan-monolaurat), pH 7.5). After washing another two times under vacuum with 200 µl each time washing-buffer 2 (10 mM Tris-HCl, 100 mM NaCl, 50 mM $MgCl_2$, pH 9.5) the wells are incubated for 5 min. with 50 µl of CSPD™ (Boehringer Mannheim, no: 1655884), diluted 1 : 100 in washing-buffer 2, which serves as a chemiluminescent substrate for the alkaline phosphatase. The solution is sucked through the membrane and after 10 min. incubation the RLU/s (Relative Light Unit per second) are detected in a Luminometer e.g. MicroLumat LB 96 P (EG&G Berthold, Wilbad, FRG).

With a serial dilution of Taq DNA polymerase a reference curve is prepared from which the linear range serves as a standard for the activity determination of the DNA polymerase to be analyzed.

Example 3

Determination of reverse transcriptase activity

The assay is performed essentially as described for determination of DNA polymerase activity except that the reaction mixture consists of the following components: 1 µg of polydA-$(dT)_{15}$, 33 µM of dTTP, 0.36 µM of digoxigenin-dUTP, 200 mg/ml BSA, 10 mM Tris-HCl, pH 8.5, 20 mM KCl, 5 mM $MgCl_2$, 10 mM DTE and various amounts of DNA polymerase The incubation temperature used is 50°C.

Example 4

Detection of DNA polymerase activity in situ

In situ PAGE analysis of polymerase activity and reverse transcriptase activity was performed essentially according to the method described by Spanos A. and Hübscher U. (1983) Methods in Enzymology 91, 263-277. Some minor, but essential modifications to the original method are, that the renaturation of the SDS-denatured polypeptides is performed in the presence of magnesium ions (3 mM) and dATP (0.5-1 µM) to assist refolding.

In brief the method is as follows:
After separation of the polypeptides from either crude cell extracts or purified samples on denaturing 8 % polyacrylamide gels (stacking gel 5 % acrylamide) which contain 150 µg activated calf thymus DNA per ml gel volume, the gels are washed four times for 30 min each at room temperature with moderate shaking in excess renaturation buffer (Tris-HCl, 50 mM, pH 8.3; EDTA, 1 mM; 2-mercaptoethanol, 3 mM; KCl, 50 mM; glycerol, 5-10 %) to remove SDS. Then the gels are incubated overnight in the same buffer, including 3 mM $MgCl_2$ and 0.5-1 µM dATP at 4°C without agitation. The first four washes are repeated the next day with renaturation buffer.
Subsequent to the removal of SDS and renaturation of the proteins the gel is transferred into reaction mixture, consisting of Tris-HCl, 50 mM, pH 8.3; KCl, 50 mM; DTT, 3 mM; $MgCl_2$, 7mM; 12 µM of each dATP, dCTP, dGTP, 8 µM dTTP and 4 µM Dig-dUTP; 10 % (vol/vol) glycerol. Gels were first incubated under shaking at room temperature for one hour and then warmed up stepwise to 37°C, 45°C, 55°C, 65°C and 72°C. At each incubation temperature DNA synthesis was allowed to proceed for 60 minutes.

After DNA synthesis, the DNA is transferred either by contact blotting or by capillary blotting (15 x SSC, Maniatis et al., supra) to nylon membranes (Boehringer Mannheim, GmbH) and crosslinked.

The detection of newly synthesized, digoxigenin-labeled DNA followed the procedure given in the previous section (Determination of DNA polymerase activity).

For molecular weight determination marker polymerases of known molecular weight (e.g. Klenow-polymerase. Pol I, Taq polymerase, Tth polymerase, HIV RT, M-MuLV RT) are applied onto the same gel, but different lanes.

The molecular weight of the claimed DNA polymerase according to this method is 100 to 105 kDa.

Example 5

Detection of 3'-5'exonuclease activity

3'-5' exonuclease activity is commonly referred as "proofreading" or "editing" activity of DNA polymerases . It is located in the small domain of the large fragment of Type A polymerases. This activity removes nucleotides from the 3' end of the primer terminus of DNA in the absence of nucleoside triphosphates (Kornberg A. and Baker T.A.(1992) DNA Replication W. H. Freemann & Company, New York). This nuclease action is suppressed by deoxynucleoside triphosphates if they match to the template and can be incorporated into the polymer.

The 3'-5' exonuclease activity of the claimed DNA polymerase can be measured as degradation or shortening of a 5'-digoxygenin-labeled oligonucleotide annealed to template DNA in the absence or presence of deoxyribonucleoside triphosphates or on DNA fragments in the absence or presence of deoxyribonucleoside triphosphates.

Degradation of digoxigenin labeled oligonucleotide: The reaction mixture is essentially the same as that for determination of DNA polymerase activity (50 mM Tris-HCl, pH 8.4, 12.5 mM $(NH_4)_2SO_4$; 10 mM KCl; 5 mM $MgCl_2$ 10 mM 2-mercaptoethanol),

except that the dNTP concentration was reduced to 12.5 µM and activated calf thymus DNA was replaced by 500 fMol primer or template/primer mixture.

The primer sequence is:

SEQ ID NO. 8.:

Dig-GCATGGATCCCACTGCCCAGGG (5' to 3'). This primer is annealed with template molecules of various 12 bp 5 prime overhangs. DNA polymerase samples of typically 0.1 units are incubated in a total reaction volume of 10 µl for 30 min at 72°C in a Perkin Elmer thermal cycler. Reactions are stopped by adding an equal volume of formamide-buffer (98 % formamide; 10 mM EDTA; bromphenol blue and xylencyanol) and denatured by heating for 10 min at 95°C. Samples are quickly chilled on ice and loaded on a 20 % denaturing polyacrylamide/urea sequencing gel. Electrophoresis is performed at 60°C and 2000 V for 2.5 hours.

After separation DNA is transferred onto a positively charged nylon membrane (Boehringer Mannheim) by contact blotting for 30 min. The DNA is crosslinked to the membrane by UV-irradiation with 120 mJoule (Stratalinker, Stratagene). The membrane is blocked with blocking solution (100 mM maleic-acid, 150 mM NaCl, 1 %(w/v) casein, pH is adjusted to 7.5 with 1 M NaOH) at room temperature for at least 30 min. digoxigenin-labeled primer DNA is detected with anti Digoxigenin-AP, Fab-fragments (Boehringer Mannheim, FRG, no 1093274) diluted 1:10000 in blocking solution (30 min at room temperature). Excess unbound antibody is removed by washing 3-4 times (10-15 min, each step) with washing buffer (100 mM maleic-acid, 150 mM NaCl, 0.3% (v/v) Tween™ 20 (Poly(oxyethylen)$_n$-sorbitan-monolaurat), pH 7.5). The membrane is transferred into a buffer containing 10 mM Tris-HCl, 100 mM NaCl, pH 9.5) and washed twice for additional 10-15 min at room temperature. Finally the membrane is soaked with a 1:1000 diluted solution of CDP-Star™ (Boehringer Mannheim). CDP-Star™ serves as a chemiluminiscent substrate for alkaline phosphatase. Then the membrane is transferred on filter paper (Whatman 3MM) to remove excess fluid, positioned between two sheets of transparent overhead foils and exposed to X-ray films (Chemiluminescent Detection Film, Boehringer Mannheim) for 5-10 min. 3'-5' exonuclease

activity is detected by degradation or shortening of the primer compared with a control (no polymerase added). As negative and positive controls DNA polymerases from Thermus aquaticus (no 3´to 5´exonuclease activity) and from Pyrococcus woeseii (exhibiting 3´to 5´exonuclease activity) are included.

Degradation of DNA fragments in the presence or absence of deoxynucleoside triphosphates: A series of dilutions of Chy polymerase was incubated for 2 hours at 70°C with 1 µg of DNA molecular weight marker III (Boehringer Mannheim) in the presence and absence of dNTPs, 1 mM each, in 50 µl of the following incuabation buffer: 50 mM Tris-HCl, pH 7.8; 10 mM $MgCl_2$; 7 mM 2-mercaptoethanol with Paraffin overlay. The DNA fragments were separated on a 1 % agarose gel containing ethidium bromide. In the absence of dNTPs a smear of DNA fragments or no DNA could be detected while in the presence of dNTPs the DNA fragments remained undegraded.

Example 6

Cloning of the Carboxydothermus hydrogenoformans DNA polymerase gene.

Preparation of chromosomal DNA from Carboxydothermus hydrogenoformans.
0.8 g biomass of Carboxydothermus hydrogenoformans was suspended in 20 ml 1 M KCl and centrifuged. Then the pellet was resuspended in 4.8 ml SET-buffer (150 mM NaCl, 15 mM EDTA, pH 8.0, 60 mM Tris-HCl, pH 8.0, 50 µg/µl RNaseA), after which 1 ml 20 % SDS and 50 µl of proteinase K (10 mg/ml) were added. The mixture was kept at 37°C for 45 minutes. After extraction with phenol and chloroform the DNA was precipitated with ethanol and dissolved in H_2O. Thus about 4.1 mg of DNA were obtained.

Amplification of specific DNA by PCR.
For amplification of the gene encoding the DNA polymerase of Carboxydothermus hydrogenoformans by the PCR technique two mixed oligonucleotides (primer 1 and 2)

were designed on the basis of conserved regions of family A DNA polymerases as published by Braithwaite D.K. and Ito J. (1993) Nucl. Acids Res. 21, 787-802.

SEQ ID No.: 1

Primer 1: 5'-CCN AAY YTN CAR AAY ATH-3'

SEQ ID No.: 2

Primer 2: 5'-YTC RTC RTG NAC YTG-3'

The PCR amplification was performed in 100 µl buffer containing 750 ng of genomic DNA from Carboxydothermus hydrogenoformans, 10 mM Tris-HCl, pH 8.8, 2.5 mM $MgCl_2$, 50 mM KCl, 200 µM dNTPs, 100 pmoles of each primer and 2.5 units of Taq polymerase (Boehringer Mannheim GmbH, FRG). The target sequence was amplified by first denaturing at 95°C for 2 min. followed by 30 cycles of 95°C for 0.5 min, 47°C for 1 min. and 72°C for 2 minutes. Thermal cycling was performed in a Perkin Elmer GenAmp 9600 thermal cycler. Agarose gel electrophoresis showed, that a fragment of approximately 600 base pairs was amplified specifically. This fragment was ligated into the pCR™II vector (Invitrogen) and the sequence determined by cycle-sequencing. The amino acid sequence deduced from this nucleotide sequence was very similar to that of other known DNA polymerases, so that primer 3 and 4 could be designed for inverse PCR.

SEQ ID No.: 3

Primer 3: 5'-GGG CGA AGA CGC TAT ATT CCT GAG C-3'

SEQ ID No.: 4

Primer 4: 5'-GAA GCC TTA ATT CAA TCT GGG AAT AAT C-3'

Inverse PCR was performed essentially as described in Triglia T. et al. (1988) Nucleic Acids Res. 16, 8186. 5 µg genomic DNA from Carboxydothermus hydrogenoformans were cleaved by EcoRI according to supplier's specifications (Boehringer Mannheim GmbH) and treated with an equal volume of phenol/chloroform mixture. The aqueous phase was removed, the DNA precipitated with ethanol and collected by centrifugation.

For circularization the digested DNA was diluted to a concentration of 50 ng/µl in ligation buffer (Boehringer Mannheim GmbH, FRG). The ligation reaction was initiated by the addition of T4 DNA Ligase (Boehringer Mannheim GmbH, FRG) to a concentration of 0.2 units/µl and the reaction was allowed to proceed for 15 hrs at 15°C. The ligated DNA was then precipitated with ethanol and collected by centrifugation.

The PCR was performed in 50 µl buffer containing 50 mM Tris-Cl, pH 9.2, 16 mM $(NH_4)_2SO_4$, 2.25 mM $MgCl_2$, 2 % (v/v) DMSO, 0.1 % (v/v) Tween™ 20, 700 ng of circularized DNA obtained as described above, 50 pmoles of each primer, 500 µM dNTP and 0.75 µl enzyme mix (Expand Long Template PCR System, Boehringer Mannheim GmbH).

The cycle conditions were as follows:

1 x denaturation of template for 2 min. at 92°C

10 x [denaturation at 92°C for 10 sec.
annealing at 64°C for 30 sec.
elongation at 68°C for 2 min.]

20 x [denaturation at 92°C for 10 sec.
annealing at 64°C for 30 sec.
elongation at 68°C for 2 min.
+ cycle elongation of 20 sec. for each cycle]

Agarose gel electrophoresis revealed a specifically amplified DNA fragment 7,000 base pairs long. The DNA fragment was ligated into the pCR™II vector (Invitrogen) and sequenced. Deduced from this sequence primer 5 and 6 coding for the 5'- and 3'-ends, respectively, of the polymerase region could be designed. Primer 5 contained a NcoI site and primer 6 contained a BamHI site.
The PCR was performed under the same conditions as described above (inverse PCR) using 750 ng genomic DNA from Carboxydothermus hydrogenoformans as template.

SEQ ID No.: 5
Primer 5: 5'-CGA ATT CAA TCC ATG GGA AAA GTA GTC CTG GTG GAT-3'

SEQ ID No.: 6
Primer 6: 5'-CGA ATT CAA GGA TCC TTA CTT CGC TTC ATA CCA GTT-3'

Cloning and expression.

The PCR product was purified by electrophoresis of 20 µl of the PCR mixture on a 0.8 % agarose gel. The 2.496 kb band of the polymerase coding region was purified from the agarose by phenol extraction. The DNA was then treated with chloroform and precipitated with ethanol. The pellet was resuspended and digested with NcoI and BamHI according to supplier's specifications (Boehringer Mannheim GmbH) to give cohesive ends for directional cloning. The DNA was ligated into the expression vector pDS56 (Stüber D., Matile H. and Garotta G. (1990) Immunological Methods. Letkovcs, I and Pernis, B., eds.) that had also been digested with NcoI and BamHI. The ligated products were introduced into E.coli strain BL21(DE3) pUBS520 (Brinkmann U. et al. (1989) Gene 85. 109-114) by transformation. Transformants were grown on L-agar

containing 100 µg/ml ampicillin and 50 µg/ml kanamycin to allow selection of recombinants. Colonies were picked and grown in L-broth containing 100 µg/ml ampicillin and 50 µg/ml kanamycin, and plasmid DNA was prepared by alkaline lysis. The plasmids were screened for insertions by digestion with BamHI. Those recombinants containing inserts were grown in L-broth containing ampicillin and kanamycin and tested for the expression of thermophilic DNA polymerase by induction of exponentially growing culture with 1 mM IPTG and assaying the heat-treated extracts for DNA polymerase activity as described above (determination of DNA polymerase activity). A recombinant expressing the DNA polymerase from Carboxydothermus hydrogenoformans was obtained. The strain was designated E.coli AR96 (DSM No. 11179) and the plasmid pAR4.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Boehringer Mannheim GmbH
(B) STREET: Sandhoferstr. 116
(C) CITY: Mannheim
(E) COUNTRY: DE
(F) POSTAL CODE (ZIP): 68305
(G) TELEPHONE: 06217595482
(H) TELEFAX: 06217594457

(ii) TITLE OF INVENTION: Thermostable DNA Polymerase from Carboxydothermus hydrogenoformans

(iii) NUMBER OF SEQUENCES: 8

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CCNAAYYTNC ARAAYATH 18

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

YTCRTCRTGN ACYTG 15

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGGCGAAGAC GCTATATTCC TGAGC 25

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GAAGCCTTAA TTCAATCTGG GAATAATC 28

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CGAATTCAAT CCATGGGAAA AGTAGTCCTG GTGGAT 36

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CGAATTCAAG GATCCTTACT TCGCTTCATA CCAGTT 36

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2496 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:1..2496

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

```
ATG GGA AAA GTA GTC CTG GTG GAT GGA AAT AGT TTA TTA CAT AGA    45
Met Gly Lys Val Val Leu Val Asp Gly Asn Ser Leu Leu His Arg
 1               5                  10                  15

GCG TTT TTT GCC CTT CCG CCC TTA AAA ACT ACT AAA GGA GAG CCT    90
Ala Phe Phe Ala Leu Pro Pro Leu Lys Thr Thr Lys Gly Glu Pro
            20                  25                  30

ACC GGG GCG GTT TAC GGG TTT TTA ACG ATG CTT TTT CGG GTA ATA   135
Thr Gly Ala Val Tyr Gly Phe Leu Thr Met Leu Phe Arg Val Ile
        35                  40                  45

AAA GAT GAA AAA CCC GAA TAT TTA GCG GTA GCT TTT GAT ATT AGC   180
Lys Asp Glu Lys Pro Glu Tyr Leu Ala Val Ala Phe Asp Ile Ser
    50                  55                  60

CGG AAA ACT TTT CGT ACC GAG CAG TTT ACT GCA TAC AAA GGG CAC   225
Arg Lys Thr Phe Arg Thr Glu Gln Phe Thr Ala Tyr Lys Gly His
65                  70                  75

CGC AAA GAA GCC CCG GAT GAG CTT GTA CCC CAG TTT GCC CTG GTG   270
Arg Lys Glu Ala Pro Asp Glu Leu Val Pro Gln Phe Ala Leu Val
                80                  85                  90

CGG GAA GTA TTA AAG GTT TTA AAT GTT CCC TAT ATT GAA CTT GAC   315
Arg Glu Val Leu Lys Val Leu Asn Val Pro Tyr Ile Glu Leu Asp
            95                  100                 105

GGT TAT GAG GCC GAT GAT ATT ATC GGC CAC CTA TCA AGG GCT TTT   360
Gly Tyr Glu Ala Asp Asp Ile Ile Gly His Leu Ser Arg Ala Phe
        110                 115                 120

GCG GGA CAA GGA CAT GAA GTG GTG ATT TAT ACC GCT GAC CGG GAC   405
Ala Gly Gln Gly His Glu Val Val Ile Tyr Thr Ala Asp Arg Asp
    125                 130                 135
```

```
ATG CTG CAA TTG GTA GAT GAA AAA ACG GTG GTA TAC CTT ACC AAA   450
Met Leu Gln Leu Val Asp Glu Lys Thr Val Val Tyr Leu Thr Lys
                140                 145                 150

AAA GGC ATT ACC GAA CTG GTT AAA ATG GAT TTA GCT GCG ATT TTA   495
Lys Gly Ile Thr Glu Leu Val Lys Met Asp Leu Ala Ala Ile Leu
                155                 160                 165

GAA AAC TAC GGC TTA AAG CCT AAA CAG CTT GTG GAT GTT AAA GGA   540
Glu Asn Tyr Gly Leu Lys Pro Lys Gln Leu Val Asp Val Lys Gly
                170                 175                 180

TTA ATG GGA GAT CCC TCG GAC AAC ATA CCC GGG GTT CCC GGG ATT   585
Leu Met Gly Asp Pro Ser Asp Asn Ile Pro Gly Val Pro Gly Ile
                185                 190                 195

GGG GAG AAA ACT GCT TTA GAT TTA ATT AAA ACT TAT GGC TCA GTG   630
Gly Glu Lys Thr Ala Leu Asp Leu Ile Lys Thr Tyr Gly Ser Val
                200                 205                 210

GAA GAA GTT TTG GCC CGT AAA GAT GAG TTA AAA CCT AAA TTA AGA   675
Glu Glu Val Leu Ala Arg Lys Asp Glu Leu Lys Pro Lys Leu Arg
                215                 220                 225

GAA AAG CTT GCC GAA CAC GAA AAT TTA GCA AAA ATA TCG AAA CAA   720
Glu Lys Leu Ala Glu His Glu Asn Leu Ala Lys Ile Ser Lys Gln
                230                 235                 240

TTA GCT ACA ATC CTG CGG GAA ATA CCG TTA GAA ATC TCC CTG GAA   765
Leu Ala Thr Ile Leu Arg Glu Ile Pro Leu Glu Ile Ser Leu Glu
                245                 250                 255

GAT TTA AAA GTT AAA GAA CCT AAT TAT GAA GAA GTT GCT AAA TTA   810
Asp Leu Lys Val Lys Glu Pro Asn Tyr Glu Glu Val Ala Lys Leu
                260                 265                 270

TTT CTT CAC CTT GAG TTT AAA AGC TTT TTA AAA GAA ATA GAA CCA   855
Phe Leu His Leu Glu Phe Lys Ser Phe Leu Lys Glu Ile Glu Pro
                275                 280                 285

AAA ATA AAG AAA GAA TAC CAG GAA GGT AAA GAT TTG GTG CAA GTT   900
Lys Ile Lys Lys Glu Tyr Gln Glu Gly Lys Asp Leu Val Gln Val
                290                 295                 300

GAA ACT GTA GAA ACG GAA GGA CAG ATT GCA GTA GTT TTT AGT GAT   945
Glu Thr Val Glu Thr Glu Gly Gln Ile Ala Val Val Phe Ser Asp
                305                 310                 315

GGA TTT TAT GTT GAT GAC GGG GAA AAA ACA AAG TTT TAC TCT TTA   990
Gly Phe Tyr Val Asp Asp Gly Glu Lys Thr Lys Phe Tyr Ser Leu
                320                 325                 330

GAC CGG CTG AAT GAA ATA GAG GAA ATA TTT AGG AAT AAA AAA ATT  1035
Asp Arg Leu Asn Glu Ile Glu Glu Ile Phe Arg Asn Lys Lys Ile
                335                 340                 345
```

```
ATT ACC GAC GAT GCC AAA GGA ATT TAT CAT GTC TGT TTA GAA AAA 1080
Ile Thr Asp Asp Ala Lys Gly Ile Tyr His Val Cys Leu Glu Lys
                350                 355                 360

GGT CTG ACT TTT CCC GAA GTT TGT TTT GAT GCG CGG ATT GCA GCT 1125
Gly Leu Thr Phe Pro Glu Val Cys Phe Asp Ala Arg Ile Ala Ala
                365                 370                 375

TAT GTT TTA AAC CCG GCC GAC CAA AAT CCC GGC CTC AAG GGG CTT 1170
Tyr Val Leu Asn Pro Ala Asp Gln Asn Pro Gly Leu Lys Gly Leu
                380                 385                 390

TAT CTA AAG TAT GAC TTA CCG GTG TAT GAA GAT GTA TCT TTA AAC 1215
Tyr Leu Lys Tyr Asp Leu Pro Val Tyr Glu Asp Val Ser Leu Asn
                395                 400                 405

ATT AGA GGG TTG TTT TAT TTA AAA AAA GAA ATG ATG AGA AAA ATC 1260
Ile Arg Gly Leu Phe Tyr Leu Lys Lys Glu Met Met Arg Lys Ile
                410                 415                 420

TTT GAG CAG GAG CAA GAA AGG TTA TTT TAT GAA ATA GAA CTT CCT 1305
Phe Glu Gln Glu Gln Glu Arg Leu Phe Tyr Glu Ile Glu Leu Pro
                425                 430                 435

TTA ACT CCA GTT CTT GCT CAA ATG GAG CAT ACC GGC ATT CAG GTT 1350
Leu Thr Pro Val Leu Ala Gln Met Glu His Thr Gly Ile Gln Val
                440                 445                 450

GAC CGG GAA GCT TTA AAA GAG ATG TCG TTA GAG CTG GGA GAG CAA 1395
Asp Arg Glu Ala Leu Lys Glu Met Ser Leu Glu Leu Gly Glu Gln
                455                 460                 465

ATT GAA GAG TTA ATC CGG GAA ATT TAT GTG CTG GCG GGG GAA GAG 1440
Ile Glu Glu Leu Ile Arg Glu Ile Tyr Val Leu Ala Gly Glu Glu
                470                 475                 480

TTT AAC TTA AAC TCG CCC AGG CAG CTG GGA GTT ATT CTT TTT GAA 1485
Phe Asn Leu Asn Ser Pro Arg Gln Leu Gly Val Ile Leu Phe Glu
                485                 490                 495

AAA CTT GGG CTG CCG GTA ATT AAA AAG ACC AAA ACG GGC TAC TCT 1530
Lys Leu Gly Leu Pro Val Ile Lys Lys Thr Lys Thr Gly Tyr Ser
                500                 505                 510

ACC GAT GCG GAG GTT TTG GAA GAG CTC TTG CCT TTC CAC GAA ATT 1575
Thr Asp Ala Glu Val Leu Glu Glu Leu Leu Pro Phe His Glu Ile
                515                 520                 525

GGC ATC GGC AAA ATA TTG AAT TAC CGG CAG CTT ATG AAG TTA AAA 1620
Ile Gly Lys Ile Leu Asn Tyr Arg Gln Leu Met Lys Leu Lys Ser
                530                 535                 540

TCC ACT TAT ACT GAC TTA ATG CCT TTA ATA AAT GAG CGT ACC GGT 1665
Thr Tyr Thr Asp Gly Leu Met Pro Leu Ile Asn Glu Arg Thr Gly
                545                 550                 555
```

```
AAA CTT CAC ACT ACT TTT AAC CAG ACC GGT ACT TTA ACC GGA CGC  1710
Lys Leu His Thr Thr Phe Asn Gln Thr Gly Thr Leu Thr Gly Arg
                560                 565                 570

CTG GCG TCT TCG GAG CCC AAT CTC CAA AAT ATT CCC ATC CGG TTG  1755
Leu Ala Ser Ser Glu Pro Asn Leu Gln Asn Ile Pro Ile Arg Leu
                575                 580                 585

GAA CTC GGT CGG AAA TTA CGC AAG ATG TTT ATA CCT TCA CCG GGG  1800
Glu Leu Gly Arg Lys Leu Arg Lys Met Phe Ile Pro Ser Pro Gly
                590                 595                 600

TAT GAT TAT ATT GTT TCG GCG GAT TAT TCC CAG ATT GAA TTA AGG  1845
Tyr Asp Tyr Ile Val Ser Ala Asp Tyr Ser Gln Ile Glu Leu Arg
                605                 610                 615

CTT CTT GCC CAT TTT TCC GAA GAG CCC AAG CTT ATT GAA GCT TAC  1890
Leu Leu Ala His Phe Ser Glu Glu Pro Lys Leu Ile Glu Ala Tyr
                620                 625                 630

CAA AAA GGG GAG GAT ATT CAC CGG AAA ACG GCC TCC GAG GTG TTC  1935
Gln Lys Gly Glu Asp Ile His Arg Lys Thr Ala Ser Glu Val Phe
                635                 640                 645

GGT GTA TCT TTG GAA GAA GTT ACT CCC GAG ATG CGC GCT CAT GCC  1980
Gly Val Ser Leu Glu Glu Val Thr Pro Glu Met Arg Ala His Ala
                650                 655                 660

AAG TCG GTG AAC TTC GGC ATT GTT TAT GGC ATT AGT GAT TTT GGT  2025
Lys Ser Val Asn Phe Gly Ile Val Tyr Gly Ile Ser Asp Phe Gly
                665                 670                 675

TTA GGC AGA GAC TTA AAG ATT CCC CGG GAG GTT GCC GGT AAG TAC  2070
Leu Gly Arg Asp Leu Lys Ile Pro Arg Glu Val Ala Gly Lys Tyr
                680                 685                 690

ATT AAA AAT TAT TTT GCC AAC TAT CCC AAA GTG CGG GAG TAT CTC  2115
Ile Lys Asn Tyr Phe Ala Asn Tyr Pro Lys Val Arg Glu Tyr Leu
                695                 700                 705

GAT GAA CTT GTC CGT ACG GCA AGA GAA AAG GGA TAT GTG ACC ACT  2160
Asp Glu Leu Val Arg Thr Ala Arg Glu Lys Gly Tyr Val Thr Thr
                710                 715                 720

TTA TTT GGG CGA AGA CGC TAT ATT CCT GAG CTA TCT TCA AAA AAC  2205
Leu Phe Gly Arg Arg Arg Tyr Ile Pro Glu Leu Ser Ser Lys Asn
                725                 730                 735

CGC ACG GTT CAG GGT TTT GGC GAA AGG ACG GCC ATG AAT ACT CCC  2250
Arg Thr Val Gln Gly Phe Gly Glu Arg Thr Ala Met Asn Thr Pro
                740                 745                 750

CTT CAG GGC TCG GCT GCC GAT ATT ATT AAG CTT GCA ATG ATT AAT  2295
Leu Gln Gly Ser Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Asn
                755                 760                 765
```

GTA GAA AAA GAA CTT AAA GCC CGT AAG CTT AAG TCC CGG CTC CTT 2340
Val Glu Lys Glu Leu Lys Ala Arg Lys Leu Lys Ser Arg Leu Leu
770 775 780

CTT TCG GTG CAC GAT GAG TTA GTT TTA GAA GTG CCG GCG GAA GAG 2385
Leu Ser Val His Asp Glu Leu Val Leu Glu Val Pro Ala Glu Glu
785 790 795

CTG GAA GAG GTA AAA GCG CTG GTA AAA GGG GTT ATG GAG TCG GTG 2430
Leu Glu Glu Val Lys Ala Leu Val Lys Gly Val Met Glu Ser Val
800 805 810

GTT GAA CTG AAA GTG CCT TTA ATC GCT GAA GTT GGT GCA GGC AAA 2475
Val Glu Leu Lys Val Pro Leu Ile Ala Glu Val Gly Ala Gly Lys
815 820 825

AAC TGG TAT GAA GCG AAG TAA
Asn Trp Tyr Glu Ala Lys *
830

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 831 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Met Gly Lys Val Val Leu Val Asp Gly Asn Ser Leu Leu His Arg
1 5 10 15

Ala Phe Phe Ala Leu Pro Pro Leu Lys Thr Thr Lys Gly Glu Pro
20 25 30

Thr Gly Ala Val Tyr Gly Phe Leu Thr Met Leu Phe Arg Val Ile
35 40 45

Lys Asp Glu Lys Pro Glu Tyr Leu Ala Val Ala Phe Asp Ile Ser
50 55 60

Arg Lys Thr Phe Arg Thr Glu Gln Phe Thr Ala Tyr Lys Gly His
65 70 75

Arg Lys Glu Ala Pro Asp Glu Leu Val Pro Gln Phe Ala Leu Val
80 85 90

Arg Glu Val Leu Lys Val Leu Asn Val Pro Tyr Ile Glu Leu Asp
95 100 105

Gly Tyr Glu Ala Asp Asp Ile Ile Gly His Leu Ser Arg Ala Phe
110 115 120

Ala Gly Gln Gly His Glu Val Val Ile Tyr Thr Ala Asp Arg Asp
125 130 135
Met Leu Gln Leu Val Asp Glu Lys Thr Val Val Tyr Leu Thr Lys
140 145 150

Lys Gly Ile Thr Glu Leu Val Lys Met Asp Leu Ala Ala Ile Leu
155 160 165

Glu Asn Tyr Gly Leu Lys Pro Lys Gln Leu Val Asp Val Lys Gly
170 175 180

Leu Met Gly Asp Pro Ser Asp Asn Ile Pro Gly Val Pro Gly Ile
185 190 195

Gly Glu Lys Thr Ala Leu Asp Leu Ile Lys Thr Tyr Gly Ser Val
200 205 210

Glu Glu Val Leu Ala Arg Lys Asp Glu Leu Lys Pro Lys Leu Arg
215 220 225

Glu Lys Leu Ala Glu His Glu Asn Leu Ala Lys Ile Ser Lys Gln
230 235 240

Leu Ala Thr Ile Leu Arg Glu Ile Pro Leu Glu Ile Ser Leu Glu
245 250 255

Asp Leu Lys Val Lys Glu Pro Asn Tyr Glu Glu Val Ala Lys Leu
260 265 270

Phe Leu His Leu Glu Phe Lys Ser Phe Leu Lys Glu Ile Glu Pro
275 280 285

Lys Ile Lys Lys Glu Tyr Gln Glu Gly Lys Asp Leu Val Gln Val
290 295 300

Glu Thr Val Glu Thr Glu Gly Gln Ile Ala Val Val Phe Ser Asp
305 310 315

Gly Phe Tyr Val Asp Asp Gly Glu Lys Thr Lys Phe Tyr Ser Leu
320 325 330

Asp Arg Leu Asn Glu Ile Glu Glu Ile Phe Arg Asn Lys Lys Ile
335 340 345

Ile Thr Asp Asp Ala Lys Gly Ile Tyr His Val Cys Leu Glu Lys
350 355 360

Gly Leu Thr Phe Pro Glu Val Cys Phe Asp Ala Arg Ile Ala Ala
365 370 375

Tyr Val Leu Asn Pro Ala Asp Gln Asn Pro Gly Leu Lys Gly Leu
380 385 390

Tyr Leu Lys Tyr Asp Leu Pro Val Tyr Glu Asp Val Ser Leu Asn
395 400 405

Ile Arg Gly Leu Phe Tyr Leu Lys Lys Glu Met Met Arg Lys Ile
410 415 420

Phe Glu Gln Glu Gln Glu Arg Leu Phe Tyr Glu Ile Glu Leu Pro
425 430 435

Leu Thr Pro Val Leu Ala Gln Met Glu His Thr Gly Ile Gln Val
440 445 450

Asp Arg Glu Ala Leu Lys Glu Met Ser Leu Glu Leu Gly Glu Gln
455 460 465

Ile Glu Glu Leu Ile Arg Glu Ile Tyr Val Leu Ala Gly Glu Glu
470 475 480

Phe Asn Leu Asn Ser Pro Arg Gln Leu Gly Val Ile Leu Phe Glu
485 490 495

Lys Leu Gly Leu Pro Val Ile Lys Lys Thr Lys Thr Gly Tyr Ser
500 505 510

Thr Asp Ala Glu Val Leu Glu Glu Leu Leu Pro Phe His Glu Ile
515 520 525

Ile Gly Lys Ile Leu Asn Tyr Arg Gln Leu Met Lys Leu Lys Ser
530 535 540

Thr Tyr Thr Asp Gly Leu Met Pro Leu Ile Asn Glu Arg Thr Gly
545 550 555

Lys Leu His Thr Thr Phe Asn Gln Thr Gly Thr Leu Thr Gly Arg
560 565 570

Leu Ala Ser Ser Glu Pro Asn Leu Gln Asn Ile Pro Ile Arg Leu
575 580 585

Glu Leu Gly Arg Lys Leu Arg Lys Met Phe Ile Pro Ser Pro Gly
590 595 600

Tyr Asp Tyr Ile Val Ser Ala Asp Tyr Ser Gln Ile Glu Leu Arg
605 610 615

Leu Leu Ala His Phe Ser Glu Glu Pro Lys Leu Ile Glu Ala Tyr
620 625 630

Gln Lys Gly Glu Asp Ile His Arg Lys Thr Ala Ser Glu Val Phe
635 640 645

Gly Val Ser Leu Glu Glu Val Thr Pro Glu Met Arg Ala His Ala
650 655 660

Lys Ser Val Asn Phe Gly Ile Val Tyr Gly Ile Ser Asp Phe Gly
665 670 675

Leu Gly Arg Asp Leu Lys Ile Pro Arg Glu Val Ala Gly Lys Tyr
680 685 690

Ile Lys Asn Tyr Phe Ala Asn Tyr Pro Lys Val Arg Glu Tyr Leu
695 700 705

Asp Glu Leu Val Arg Thr Ala Arg Glu Lys Gly Tyr Val Thr Thr
710 715 720

Leu Phe Gly Arg Arg Arg Tyr Ile Pro Glu Leu Ser Ser Lys Asn
725 730 735

Arg Thr Val Gln Gly Phe Gly Glu Arg Thr Ala Met Asn Thr Pro
740 745 750

Leu Gln Gly Ser Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Asn
755 760 765

Val Glu Lys Glu Leu Lys Ala Arg Lys Leu Lys Ser Arg Leu Leu
770 775 780

Leu Ser Val His Asp Glu Leu Val Leu Glu Val Pro Ala Glu Glu
785 790 795

Leu Glu Glu Val Lys Ala Leu Val Lys Gly Val Met Glu Ser Val
800 805 810

Val Glu Leu Lys Val Pro Leu Ile Ala Glu Val Gly Ala Gly Lys
815 820 825

Asn Trp Tyr Glu Ala Lys *
830

INTERNATIONAL FORM

Boehringer Mannheim GmbH

Sandhofer Str. 116

68305 Mannheim

RECEIPT IN THE CASE OF AN ORIGINAL DEPOSIT
issued pursuant to Rule 7.1 by the
INTERNATIONAL DEPOSITARY AUTHORITY
identified at the bottom of this page

I. IDENTIFICATION OF THE MICROORGANISM	
Identification reference given by the DEPOSITOR: DSM 8979	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: DSM 8979

II. SCIENTIFIC DESCRIPTION AND/OR PROPOSED TAXONOMIC DESIGNATION

The microorganism identified under I. above was accompanied by:

( ) a scientific description
(X) a proposed taxonomic designation

(Mark with a cross where applicable).

III. RECEIPT AND ACCEPTANCE

This International Depositary Authority accepts the microorganism identified under I. above, which was received by it on 1994-02-10 (Date of the original deposit)[1].

IV. RECEIPT OF REQUEST FOR CONVERSION

The microorganism identified under I above was received by this International Depositary Authority on (date of original deposit) and a request to convert the original deposit to a deposit under the Budapest Treaty was received by it on (date of receipt of request for conversion).

V. INTERNATIONAL DEPOSITARY AUTHORITY

Name: DSM-DEUTSCHE SAMMLUNG VON MIKROORGANISMEN UND ZELLKULTUREN GmbH Address: Mascheroder Weg 1b D-38124 Braunschweig	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s): U Weirs Date: 1994-09-21

[1] Where Rule 6.4 (d) applies, such date is the date on which the status of international depositary authority was acquired.

Form DSM-BP/4 (sole page) 07/94

INTERNATIONAL FORM

Boehringer Mannheim GmbH

Sandhofer Str. 116

68305 Mannheim

VIABILITY STATEMENT
issued pursuant to Rule 10.2 by the
INTERNATIONAL DEPOSITARY AUTHORITY
identified at the bottom of this page

I. DEPOSITOR	II. IDENTIFICATION OF THE MICROORGANISM
Name: Boehringer Mannheim GmbH Address: Sandhofer Str. 116 68305 Mannheim	Accession number given by the INTERNATIONAL DEPOSITARY AUTHORITY: DSM 8979 Date of the deposit or the transfer[1]: 1994-02-10

III. VIABILITY STATEMENT

The viability of the microorganism identified under II above was tested on 1994-02-16 [2].
On that date, the said microorganism was

(X)[3] viable

( )[3] no longer viable

IV. CONDITIONS UNDER WHICH THE VIABILITY TEST HAS BEEN PERFORMED[4]

V. INTERNATIONAL DEPOSITARY AUTHORITY

Name: DSM-DEUTSCHE SAMMLUNG VON MIKROORGANISMEN UND ZELLKULTUREN GmbH Address: Mascheroder Weg 1b D-38124 Braunschweig	Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s): J. Weiss Date: 1994-09-21

[1] Indicate the date of original deposit or, where a new deposit or a transfer has been made, the most recent relevant date (date of the new deposit or date of the transfer).
[2] In the cases referred to in Rule 10.2(a) (ii) and(iii), refer to the most recent viability test.
[3] Mark with a cross the applicable box.
[4] Fill in if the information has been requested and if the results of the test were negative.

Form DSM-BP/9 (sole page) 07/94

Claims

1. A purified thermostable DNA polymerase obtainable from Carboxydothermus hydrogenoformans that catalyses the template directed polymerisation of DNA, posesses 3'-5' exonuclease activity and has reverse transcriptase activity in the presence of magnesiums ions and in the substantial absence of manganese ions.

2. The polymerase according to claim 1, wherein said polymerase exhibits a reverse transcriptase activity which is Mn^{2+} dependent.

3. The polymerase as claimed in any of claims 1-2, wherein the magnesium ions dependent reverse transcriptase activity of said polymerase is higher than the DNA polymerase activity of said polymerase.

4. The polymerase as claimed in any of claims 1-3, wherein the magnesium ions dependent reverse transcriptase activity of said polymerase is higher than the manganese dependent reverse transcriptase activity of said polymerase.

5. The polymerase as claimed in any of claims 1-4, wherein said polymerase has an apparent molecular weight between about 100 to 105 kDa as determined by SDS polyacrylamide electrophoresis.

6. The polymerase as claimed in any of claims 1-5, wherein said polymerase is obtainable from E. coli BL21 (DE3) pUBS520, the strain being designated AR96.

7. An isolated DNA sequence coding for the polymerase as claimed in any one of claims 1-6 obtainable from Carboxydothermus hydrogenoformans.

8. A recombinant DNA sequence capable of encoding polymerase activity of the microorganism Carboxydothermus hydrogenoformans.

9. An isolated DNA sequence represented by the formula shown in SEQ ID No. 7.

10. A vector containing the isolated DNA sequence as claimed in any of claims 7-9.

11. The vector according to claim 10, wherein such vector is plasmid pDS56 carrying the Carboxydothermus hydrogenoformas DNA polymerase gene and is then designated pAR4.

12. The vector according to claims 10 and 11 providing some or all of the following features:

(1) promotors or sites of initiation of transcription
(2) operators which could be used to turn gene expression on or off
(3) ribosome binding sites for improved translation
(4) transcription or translation termination sites

13. A microbial host transformed with the vector of claims 10-12.

14. A microbial host according to claim 13 wherein said transformant is from E. coli BL21 (DE3) pUBS520, the strain being designated AR96.

15. A process for the preparation of DNA polymerase according to any of the claims 1-6 comprising the steps:

(a) culturing the natural strain Carboxydothermus hydrogenoformas
(b) suspending the cells of the natural strain in buffer
(c) disrupting the cells
(d) purifying the DNA polymerase by chromatographic steps including the use of one or more Sepharose-columns.

16. A process for the preparation of DNA polymerase according to any one of claims 1-6 comprising growing a recombinant E. coli strain transformed with a vector according to claims 10-12 and purifying and isolating the DNA polymerase.

17. A process of amplifying DNA, characterized in that a thermostable DNA polymerase as claimed in any one of claims 1-6 is used.

18. A process for second cDNA cloning and DNA sequencing, characterized in that a thermostable DNA polymerase as claimed in any one of claims 1-6 is used.

19. A process for DNA labeling, characterized in that a thermostable DNA polymerase as claimed in any one of claims 1-6 is used.

20. A process for reverse transcription. characterized in that a thermostable DNA polymerase as claimed in any one of claims 1-6 is used.

Figure 1:

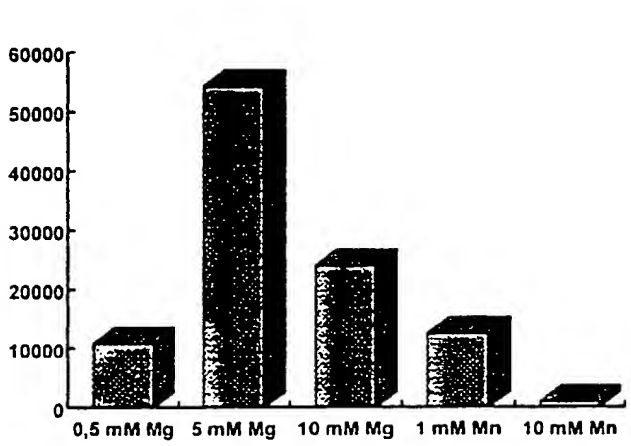
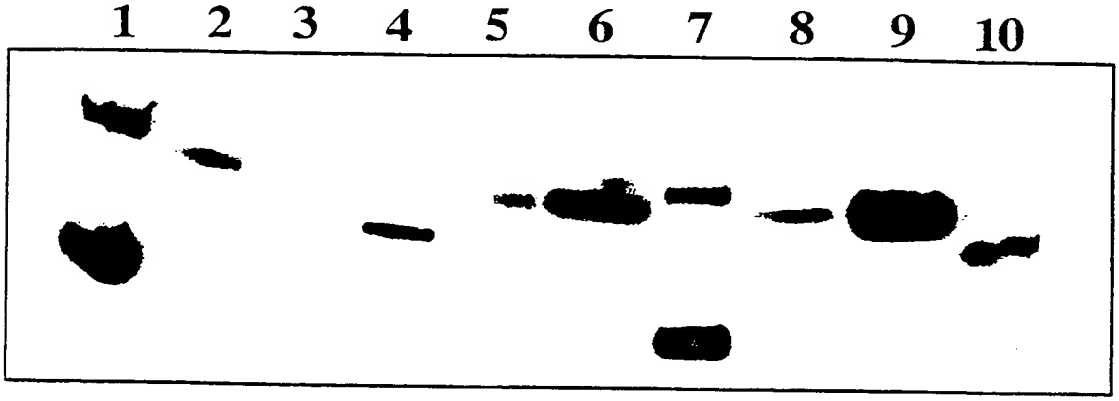

RLU
60000
50000
40000
30000
20000
10000
0
0,5 mM Mg
5 mM Mg
10 mM Mg
1 mM Mn
10 mM Mn

Figure 2:

Detection of reverse transcriptase activity of DNA polymerase from *Carboxydothermus hydrogenoformans* and reference polymerases *in situ*.



1: E.coli pol I (20U) + Klenow fragment (10U)
2: Cell extract from C. hydrogenoformans (2µl)
3: Chy polymerase, rec. (7U)
4: Taq polymerase (20U)
5: Cell extract from C. hydrogenoformans (2µl)
6: Chy polymerase, rec. (14U)
7: E.coli pol I (10U) + Klenow fragment (5U)
8: Cell extract from C. hydrogenoformans (4µl)
9: Chy polymerase, rec. (70U)
10: Taq polymerase (20U)

Figure 3:

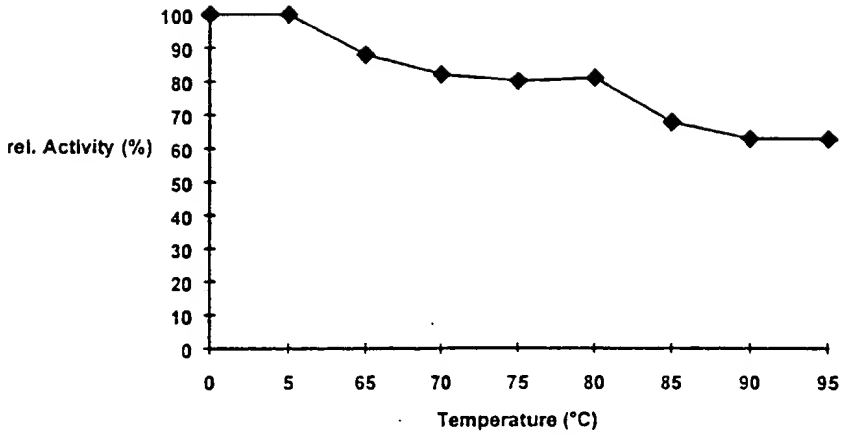

rel. Activity (%)
100
90
80
70
60
50
40
30
20
10
0
0
5
65
70
75
80
85
90
95
Temperature (°C)

Fig.4

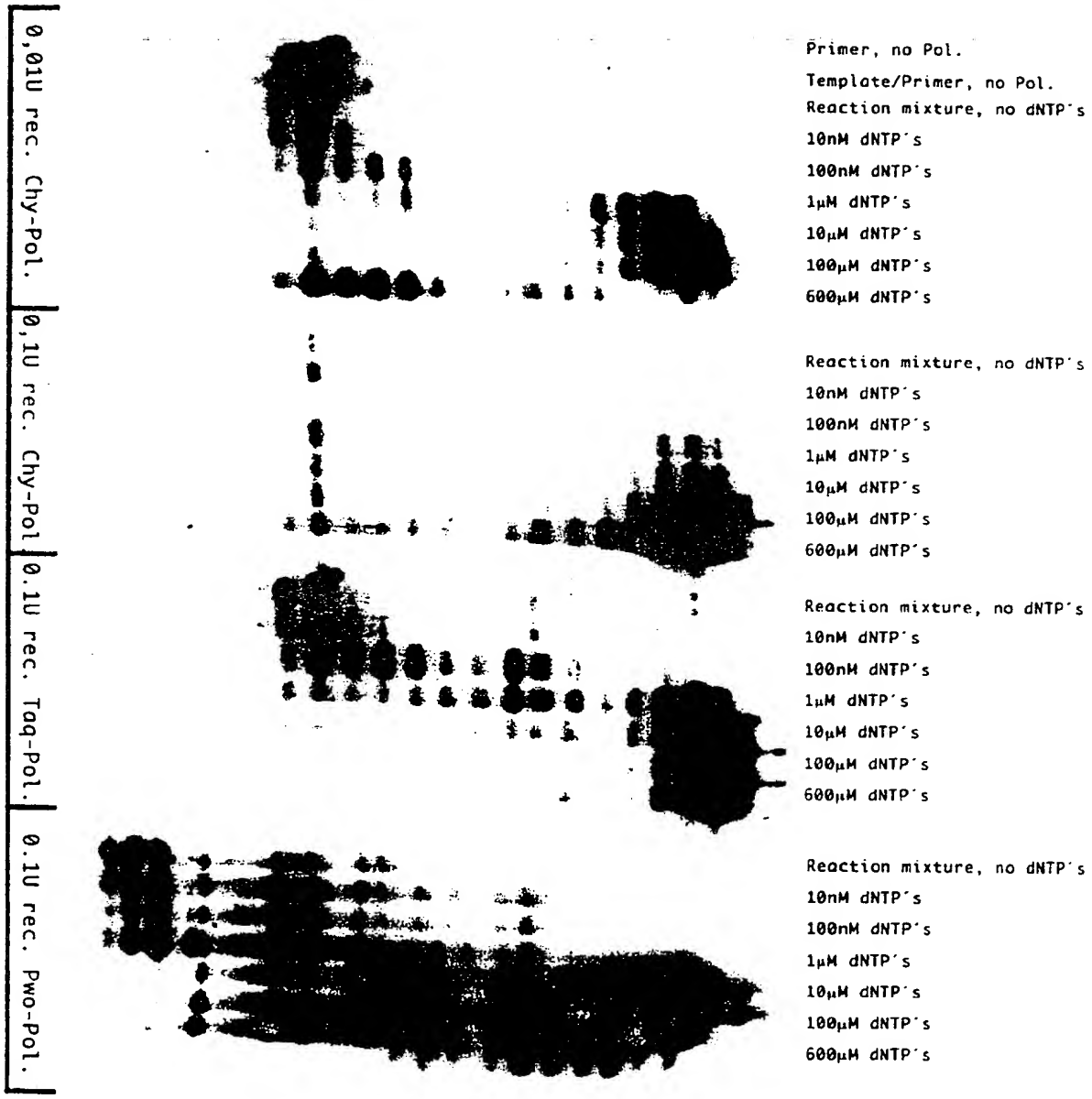

0,01U rec. Chy-Pol.
0,1U rec. Chy-Pol
0.1U rec. Taq-Pol.
0.1U rec. Pwo-Pol.
Primer, no Pol.
Template/Primer, no Pol.
Reaction mixture, no dNTP's
10nM dNTP's
100nM dNTP's
1µM dNTP's
10µM dNTP's
100µM dNTP's
600µM dNTP's
Reaction mixture, no dNTP's
10nM dNTP's
100nM dNTP's
1µM dNTP's
10µM dNTP's
100µM dNTP's
600µM dNTP's
Reaction mixture, no dNTP's
10nM dNTP's
100nM dNTP's
1µM dNTP's
10µM dNTP's
100µM dNTP's
600µM dNTP's
Reaction mixture, no dNTP's
10nM dNTP's
100nM dNTP's
1µM dNTP's
10µM dNTP's
100µM dNTP's
600µM dNTP's

Figure 5:

SEQ ID No.: 7
SEQ ID NO.: 8

```
ATG GGA AAA GTA GTC CTG GTG GAT GGA AAT AGT TTA TTA CAT AGA GCG  48
 M   G   K   V   V   L   V   D   G   N   S   L   L   H   R   A  16
TTT TTT GCC CTT CCG CCC TTA AAA ACT ACT AAA GGA GAG CCT ACC GGG  96
 F   F   A   L   P   P   L   K   T   T   K   G   E   P   T   G  32
GCG GTT TAC GGG TTT TTA ACG ATG CTT TTT CGG GTA ATA AAA GAT GAA 144
 A   V   Y   G   F   L   T   M   L   F   R   V   I   K   D   E  48
AAA CCC GAA TAT TTA GCG GTA GCT TTT GAT ATT AGC CGG AAA ACT TTT 192
 K   P   E   Y   L   A   V   A   F   D   I   S   R   K   T   F  64
CGT ACC GAG CAG TTT ACT GCA TAC AAA GGG CAC CGC AAA GAA GCC CCG 240
 R   T   E   Q   F   T   A   Y   K   G   H   R   K   E   A   P  80
GAT GAG CTT GTA CCC CAG TTT GCC CTG GTG CGG GAA GTA TTA AAG GTT 288
 D   E   L   V   P   Q   F   A   L   V   R   E   V   L   K   V  96
TTA AAT GTT CCC TAT ATT GAA CTT GAC GGT TAT GAG GCC GAT GAT ATT 336
 L   N   V   P   Y   I   E   L   D   G   Y   E   A   D   D   I 112
```

```
ATC GGC CAC CTA TCA AGG GCT TTT GCG GGA CAA GGA CAT GAA GTG GTG  384
 I   G   H   L   S   R   A   F   A   G   Q   G   H   E   V   V  128
ATT TAT ACC GCT GAC CGG GAC ATG CTG CAA TTG GTA GAT GAA AAA ACG  432
 I   Y   T   A   D   R   D   M   L   Q   L   V   D   E   K   T  144
GTG GTA TAC CTT ACC AAA AAA GGC ATT ACC GAA CTG GTT AAA ATG GAT  480
 V   V   Y   L   T   K   K   G   I   T   E   L   V   K   M   D  160
TTA GCT GCG ATT TTA GAA AAC TAC GGC TTA AAG CCT AAA CAG CTT GTG  528
 L   A   A   I   L   E   N   Y   G   L   K   P   K   Q   L   V  176
GAT GTT AAA GGA TTA ATG GGA GAT CCC TCG GAC AAC ATA CCC GGG GTT  576
 D   V   K   G   L   M   G   D   P   S   D   N   I   P   G   V  192
CCC GGG ATT GGG GAG AAA ACT GCT TTA GAT TTA ATT AAA ACT TAT GGC  624
 P   G   I   G   E   K   T   A   L   D   L   I   K   T   Y   G  208
TCA GTG GAA GAA GTT TTG GCC CGT AAA GAT GAG TTA AAA CCT AAA TTA  672
 S   V   E   E   V   L   A   R   K   D   E   L   K   P   K   L  224
AGA GAA AAG CTT GCC GAA CAC GAA AAT TTA GCA AAA ATA TCG AAA CAA  720
 R   E   K   L   A   E   H   E   N   L   A   K   I   S   K   Q  240
```

```
TTA GCT ACA ATC CTG CGG GAA ATA CCG TTA GAA ATC TCC CTG GAA GAT  768
 L   A   T   I   L   R   E   I   P   L   E   I   S   L   E   D  256
TTA AAA GTT AAA GAA CCT AAT TAT GAA GAA GTT GCT AAA TTA TTT CTT  816
 L   K   V   K   E   P   N   Y   E   E   V   A   K   L   F   L  272
CAC CTT GAG TTT AAA AGC TTT TTA AAA GAA ATA GAA CCA AAA ATA AAG  864
 H   L   E   F   K   S   F   L   K   E   I   E   P   K   I   K  288
AAA GAA TAC CAG GAA GGT AAA GAT TTG GTG CAA GTT GAA ACT GTA GAA  912
 K   E   Y   Q   E   G   K   D   L   V   Q   V   E   T   V   E  304
ACG GAA GGA CAG ATT GCA GTA GTT TTT AGT GAT GGA TTT TAT GTT GAT  960
 T   E   G   Q   I   A   V   V   F   S   D   G   F   Y   V   D  320
GAC GGG GAA AAA ACA AAG TTT TAC TCT TTA GAC CGG CTG AAT GAA ATA 1008
 D   G   E   K   T   K   F   Y   S   L   D   R   L   N   E   I  336
GAG GAA ATA TTT AGG AAT AAA AAA ATT ATT ACC GAC GAT GCC AAA GGA 1056
 E   E   I   F   R   N   K   K   I   I   T   D   D   A   K   G  352
ATT TAT CAT GTC TGT TTA GAA AAA GGT CTG ACT TTT CCC GAA GTT TGT 1104
 I   Y   H   V   C   L   E   K   G   L   T   F   P   E   V   C  368
```

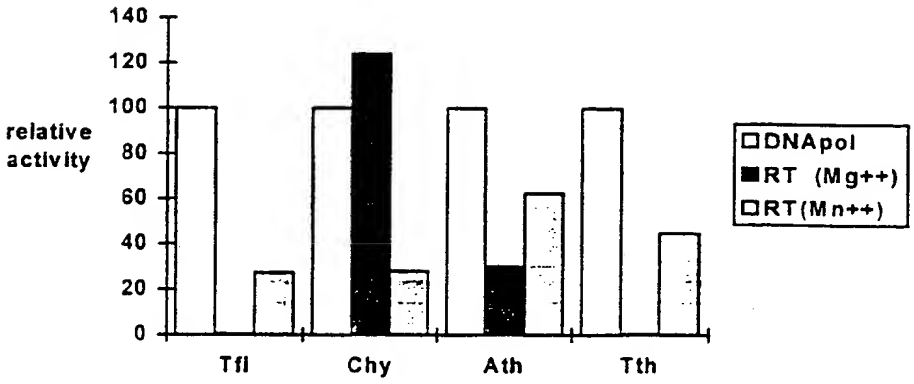
```
TTT GAT GCG CGG ATT GCA GCT TAT GTT TTA AAC CCG GCC GAC CAA AAT  1152
 F   D   A   R   I   A   A   Y   V   L   N   P   A   D   Q   N   384
CCC GGC CTC AAG GGG CTT TAT CTA AAG TAT GAC TTA CCG GTG TAT GAA  1200
 P   G   L   K   G   L   Y   L   K   Y   D   L   P   V   Y   E   400
GAT GTA TCT TTA AAC ATT AGA GGG TTG TTT TAT TTA AAA AAA GAA ATG  1248
 D   V   S   L   N   I   R   G   L   F   Y   L   K   K   E   M   416
ATG AGA AAA ATC TTT GAG CAG GAG CAA GAA AGG TTA TTT TAT GAA ATA  1296
 M   R   K   I   F   E   Q   E   Q   E   R   L   F   Y   E   I   432
GAA CTT CCT TTA ACT CCA GTT CTT GCT CAA ATG GAG CAT ACC GGC ATT  1344
 E   L   P   L   T   P   V   L   A   Q   M   E   H   T   G   I   448
CAG GTT GAC CGG GAA GCT TTA AAA GAG ATG TCG TTA GAG CTG GGA GAG  1392
 Q   V   D   R   E   A   L   K   E   M   S   L   E   L   G   E   464
CAA ATT GAA GAG TTA ATC CGG GAA ATT TAT GTG CTG GCG GGG GAA GAG  1440
 Q   I   E   E   L   I   R   E   I   Y   V   L   A   G   E   E   480
TTT AAC TTA AAC TCG CCC AGG CAG CTG GGA GTT ATT CTT TTT GAA AAA  1488
 F   N   L   N   S   P   R   Q   L   G   V   I   L   F   E   K   496
```

```
CTT GGG CTG CCG GTA ATT AAA AAG ACC AAA ACG GGC TAC TCT ACC GAT  1536
 L   G   L   P   V   I   K   K   T   K   T   G   Y   S   T   D  512
GCG GAG GTT TTG GAA GAG CTC TTG CCT TTC CAC GAA ATT ATC GGC AAA  1584
 A   E   V   L   E   E   L   L   P   F   H   E   I   I   G   K  528
ATA TTG AAT TAC CGG CAG CTT ATG AAG TTA AAA TCC ACT TAT ACT GAC  1632
 I   L   N   Y   R   Q   L   M   K   L   K   S   T   Y   T   D  544
GGC TTA ATG CCT TTA ATA AAT GAG CGT ACC GGT AAA CTT CAC ACT ACT  1680
 G   L   M   P   L   I   N   E   R   T   G   K   L   H   T   T  560
TTT AAC CAG ACC GGT ACT TTA ACC GGA CGC CTG GCG TCT TCG GAG CCC  1728
 F   N   Q   T   G   T   L   T   G   R   L   A   S   S   E   P  576
AAT CTC CAA AAT ATT CCC ATC CGG TTG GAA CTC GGT CGG AAA TTA CGC  1776
 N   L   Q   N   I   P   I   R   L   E   L   G   R   K   L   R  592
AAG ATG TTT ATA CCT TCA CCG GGG TAT GAT TAT ATT GTT TCG GCG GAT  1824
 K   M   F   I   P   S   P   G   Y   D   Y   I   V   S   A   D  608
TAT TCC CAG ATT GAA TTA AGG CTT CTT GCC CAT TTT TCC GAA GAG CCC  1872
 Y   S   Q   I   E   L   R   L   L   A   H   F   S   E   E   P  624
```

```
AAG CTT ATT GAA GCT TAC CAA AAA GGG GAG GAT ATT CAC CGG AAA ACG  1920
 K   L   I   E   A   Y   Q   K   G   E   D   I   H   R   K   T   640
GCC TCC GAG GTG TTC GGT GTA TCT TTG GAA GAA GTT ACT CCC GAG ATG  1968
 A   S   E   V   F   G   V   S   L   E   E   V   T   P   E   M   656
CGC GCT CAT GCC AAG TCG GTG AAC TTC GGC ATT GTT TAT GGC ATT AGT  2016
 R   A   H   A   K   S   V   N   F   G   I   V   Y   G   I   S   672
GAT TTT GGT TTA GGC AGA GAC TTA AAG ATT CCC CGG GAG GTT GCC GGT  2064
 D   F   G   L   G   R   D   L   K   I   P   R   E   V   A   G   688
AAG TAC ATT AAA AAT TAT TTT GCC AAC TAT CCC AAA GTG CGG GAG TAT  2112
 K   Y   I   K   N   Y   F   A   N   Y   P   K   V   R   E   Y   704
CTC GAT GAA CTT GTC CGT ACG GCA AGA GAA AAG GGA TAT GTG ACC ACT  2160
 L   D   E   L   V   R   T   A   R   E   K   G   Y   V   T   T   720
TTA TTT GGG CGA AGA CGC TAT ATT CCT GAG CTA TCT TCA AAA AAC CGC  2208
 L   F   G   R   R   R   Y   I   P   E   L   S   S   K   N   R   736
ACG GTT CAG GGT TTT GGC GAA AGG ACG GCC ATG AAT ACT CCC CTT CAG  2256
 T   V   Q   G   F   G   E   R   T   A   M   N   T   P   L   Q   752
```

```
GGC TCG GCT GCC GAT ATT ATT AAG CTT GCA ATG ATT AAT GTA GAA AAA  2304
 G   S   A   A   D   I   I   K   L   A   M   I   N   V   E   K   768
GAA CTT AAA GCC CGT AAG CTT AAG TCC CGG CTC CTT CTT TCG GTG CAC  2352
 E   L   K   A   R   K   L   K   S   R   L   L   L   S   V   H   784
GAT GAG TTA GTT TTA GAA GTG CCG GCG GAA GAG CTG GAA GAG GTA AAA  2400
 D   E   L   V   L   E   V   P   A   E   E   L   E   E   V   K   800
GCG CTG GTA AAA GGG GTT ATG GAG TCG GTG GTT GAA CTG AAA GTG CCT  2448
 A   L   V   K   G   V   M   E   S   V   V   E   L   K   V   P   816
TTA ATC GCT GAA GTT GGT GCA GGC AAA AAC TGG TAT GAA GCG AAG TAA  2496
 L   I   A   E   V   G   A   G   K   N   W   Y   E   A   K   *   831
```

Figure 6:


140
120
100
80
60
40
20
0
relative activity
Tfl
Chy
Ath
Tth
DNApol
RT (Mg++)
RT(Mn++)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 97/05391

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/52 C12N9/12 C12N15/70 C12N1/21 C12P21/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 10640 A (LIFE TECHNOLOGIES, INC.) 11 April 1996 see page 12, line 1 - page 19, line 31; examples 7,8 ---	1,2,5,7, 8,10,12, 13,16-20
X	WO 92 03556 A (CETUS CORPORATION) 5 March 1992 see page 2, line 28 - page 3, line 16 see page 14, line 32 - page 17, line 12 see page 32, line 11 - line 37; examples 4,5 --- -/--	1,2,5,7, 8,10,12, 13,16-20

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 January 1998	13. 02. 98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Donath, C

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PERLER, F.B. ET AL.: "Thermostable DNA Polymerases" ADVANCES IN PROTEIN CHEMISTRY, RICHARDS,F.M. ET AL. (ED.), ACADEMIC PRESS LONDON, vol. 48, 1996, pages 377-435, XP000654858 see the whole document	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter. ional [redacted] tion No
PCT/EP 97/05391

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9610640 A	11-04-96	CA 2174944 A	11-04-96
		EP 0725827 A	14-08-96
		JP 9506783 T	08-07-97
WO 9203556 A	05-03-92	US 5374553 A	20-12-94
		AU 653747 B	13-10-94
		AU 8501491 A	17-03-92
		CA 2089495 A	14-02-92
		DE 544789 T	16-12-93
		EP 0544789 A	09-06-93
		JP 8298991 A	19-11-96
		JP 2584198 B	19-02-97
		JP 7147990 A	13-06-95
		JP 7108220 B	22-11-95
		JP 6500020 T	06-01-94
		US 5624833 A	29-04-97
		US 5420029 A	30-05-95